Computer Vision

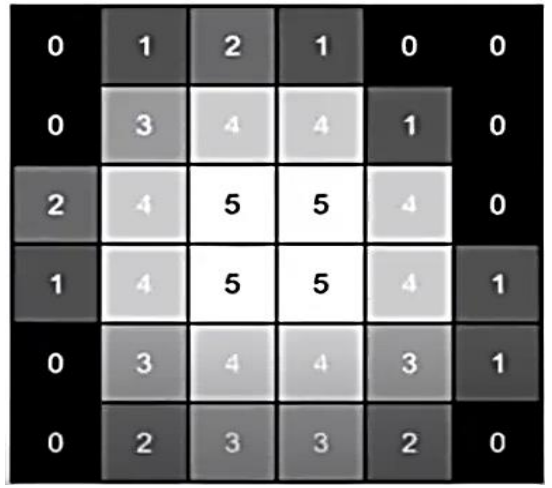
Section 5

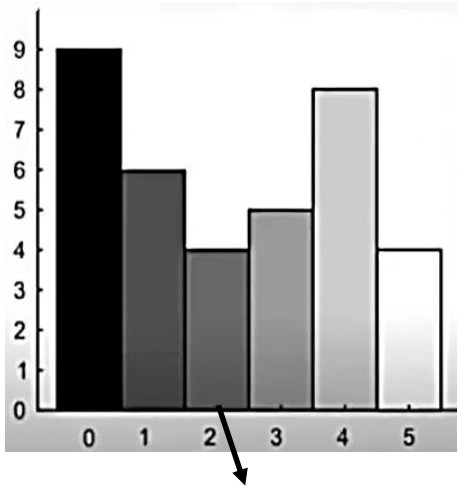
Image Segmentation

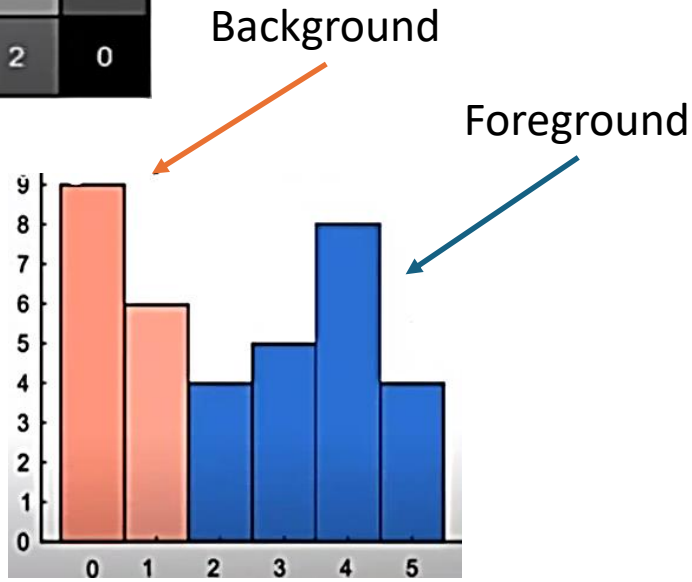
- Image segmentation is a computer vision task that involves dividing an image into multiple meaningful parts or regions to make it easier to analyze.
- Semantic Segmentation
  - Classifies each pixel into a category but doesn't distinguish between instances.
- Instance Segmentation
  - Similar to semantic segmentation but also differentiates between separate instances of the same object class.

Thresholding (Otsu Method)

- Thresholding is a technique to convert a grayscale image into a binary image (black and white) by turning all pixels above or below a certain intensity value to white (255) or black (0).
- Otsu's method automatically finds an optimal threshold value to separate an image into foreground and background based on its histogram.







Thresholding (Otsu Method)

- We try every threshold t and compute:

$$\sigma\_b^2(t) = \omega_1 * \omega_2 * (\mu_1 - \mu_2)^2$$

  - $\omega_1(t)$ = probability of class 1 (≤ t)
  - $\omega_2(t)$ = probability of class 2 (> t)
  - $\mu_1(t)$ = mean of class 1
  - $\mu_2(t)$ = mean of class 2

Threshold t = 2

Class 1: [0, 1, 2] → 9+6+4 = 19 pixels

- $\omega_1 = 19 / 36 \approx 0.528$
- $\mu_1 = (0\times9 + 1\times6 + 2\times4) / 19 = (0 + 6 + 8) / 19 \approx$ **0.7368**

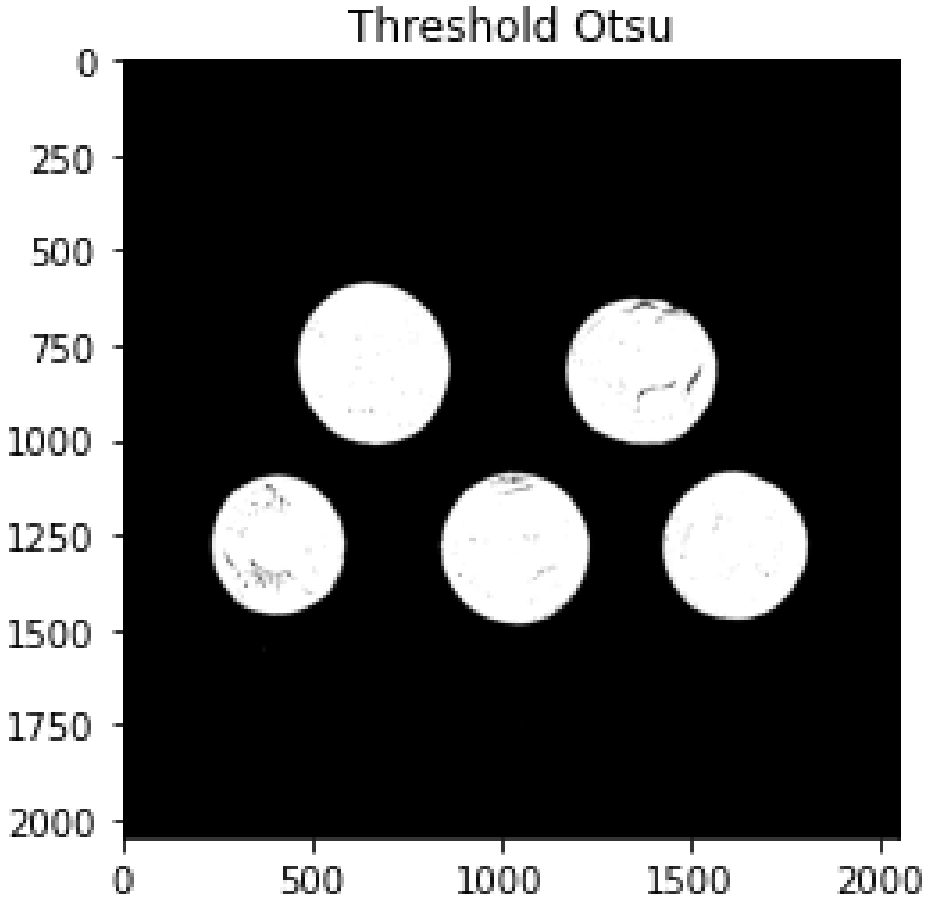
Class 2: [3, 4, 5] → 17 pixels

- $\omega_2 = 17 / 36 \approx 0.472$
- $\mu_2 = (3\times5 + 4\times8 + 5\times4) / 17 = (15 + 32 + 20) / 17 \approx$ **3.7059**

$\sigma\_b^2 =$ **0.528 × 0.472 × (0.7368 - 3.7059)²**

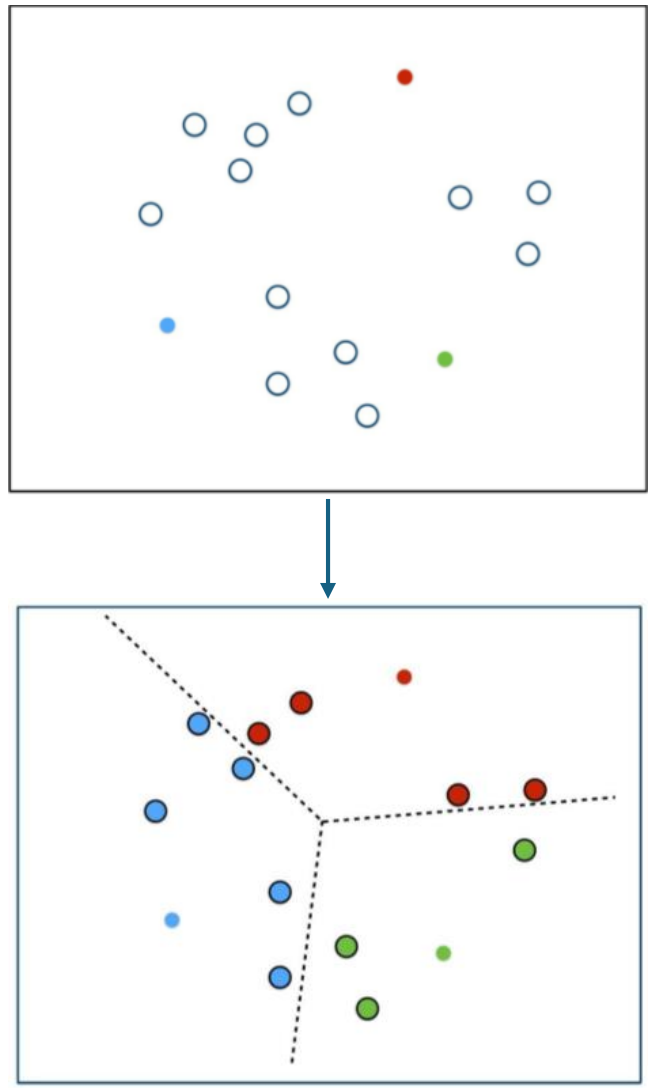
$\approx 0.528 \times 0.472 \times 8.819 \approx$ **2.196**

Thresholding (Otsu Method)

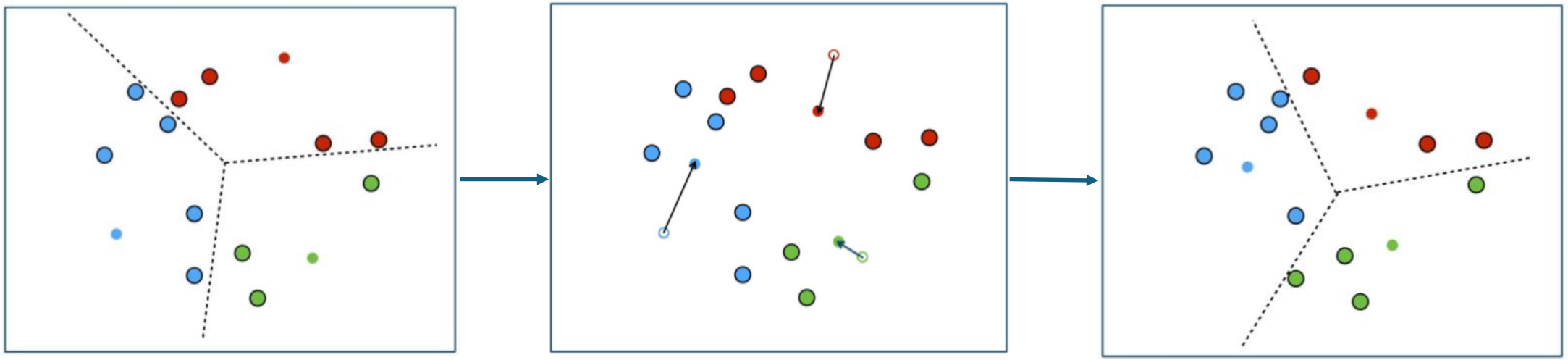


K-means Clustering

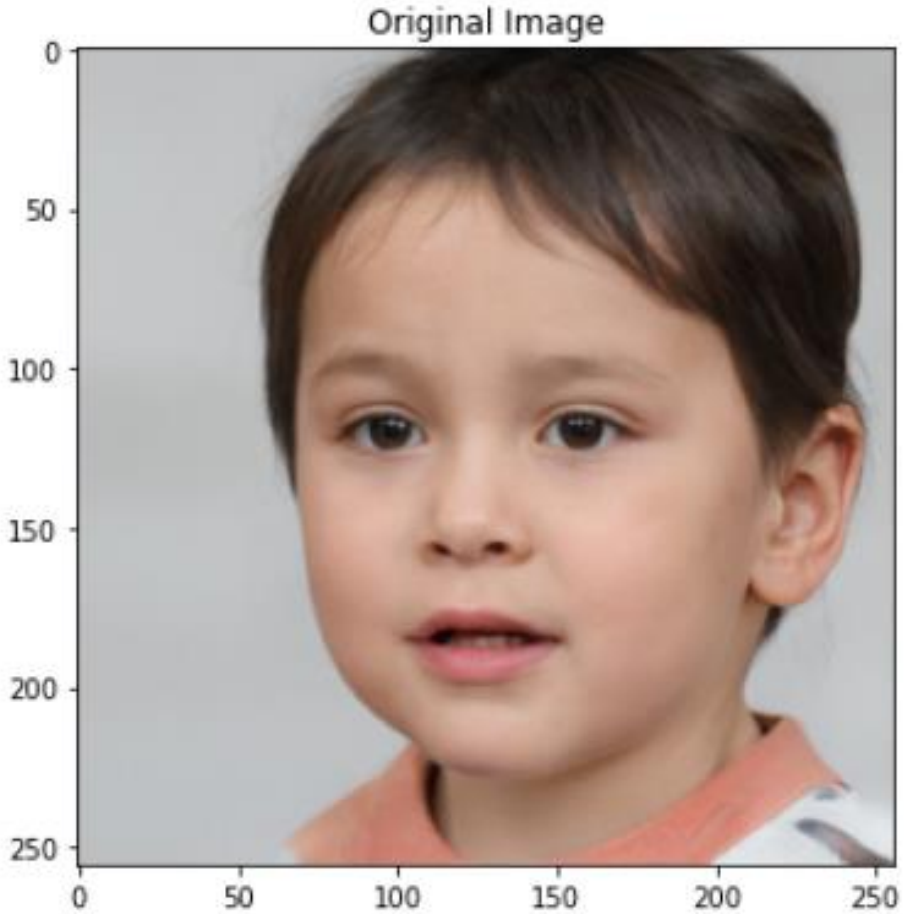
- K-means is an unsupervised machine learning algorithm that groups similar data points into K clusters.
- Step by Step:
  - Choose K – the number of clusters you want (e.g., K = 3)
  - Initialize K centroids randomly
  - Assign each data point to the nearest centroid
  - Recalculate the centroids by averaging the points assigned to each
  - Repeat steps 3-4 until centroids don’t change much (or max iterations)

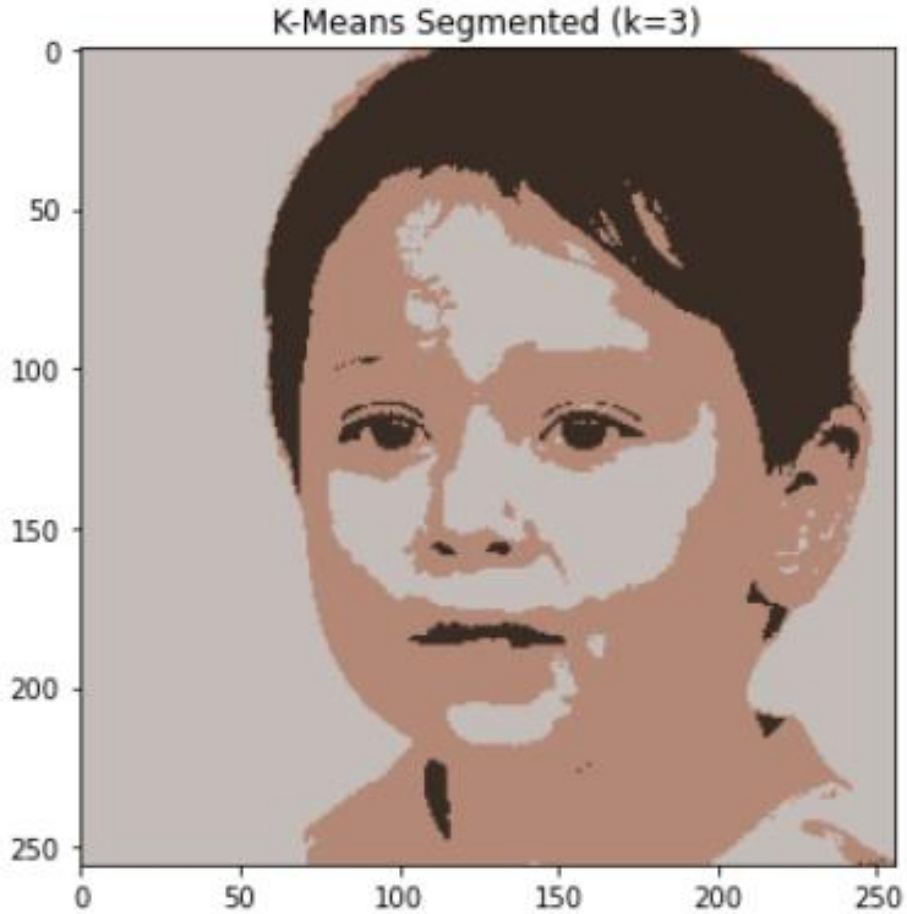


K-means Clustering



K-means Clustering





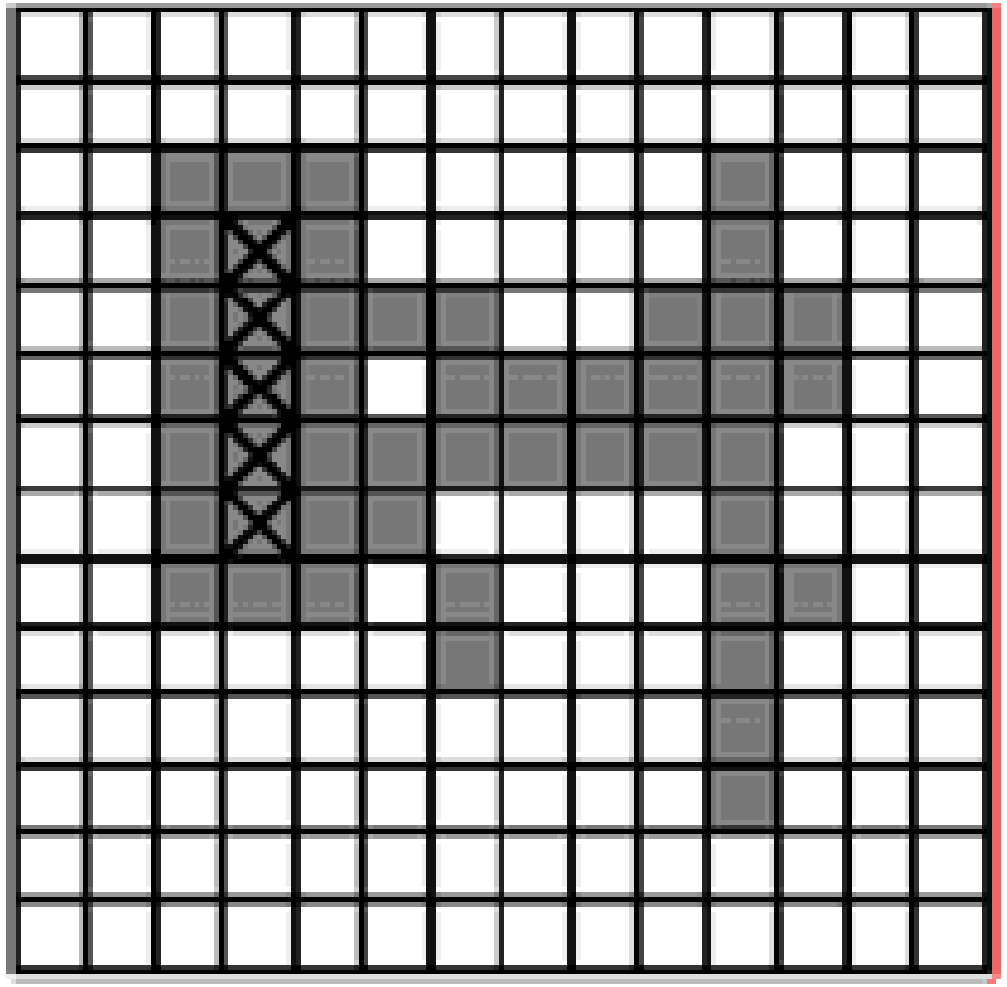
Morphological operations

- Morphological operations are image processing techniques that focus on the shape and structure of objects in an image.
- They are most commonly applied to binary images (black & white).
- To **clean up**, **refine**, or **analyze** shapes in an image:
  - Remove noise (tiny specks)
  - Fill holes inside objects
  - Separate or connect objects
  - Highlight outlines

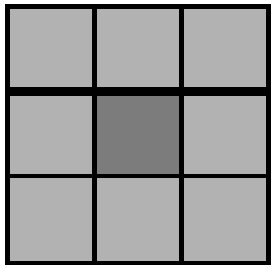
Morphological operations

Operation	What It Does	Visual
Erosion	Shrinks white regions (foreground)	Erodes boundaries
Dilation	Grows white regions	Expands object size
Opening	Erosion → Dilation	Removes small noise
Closing	Dilation → Erosion	Fills small holes
Gradient	Dilation - Erosion	Extracts edges

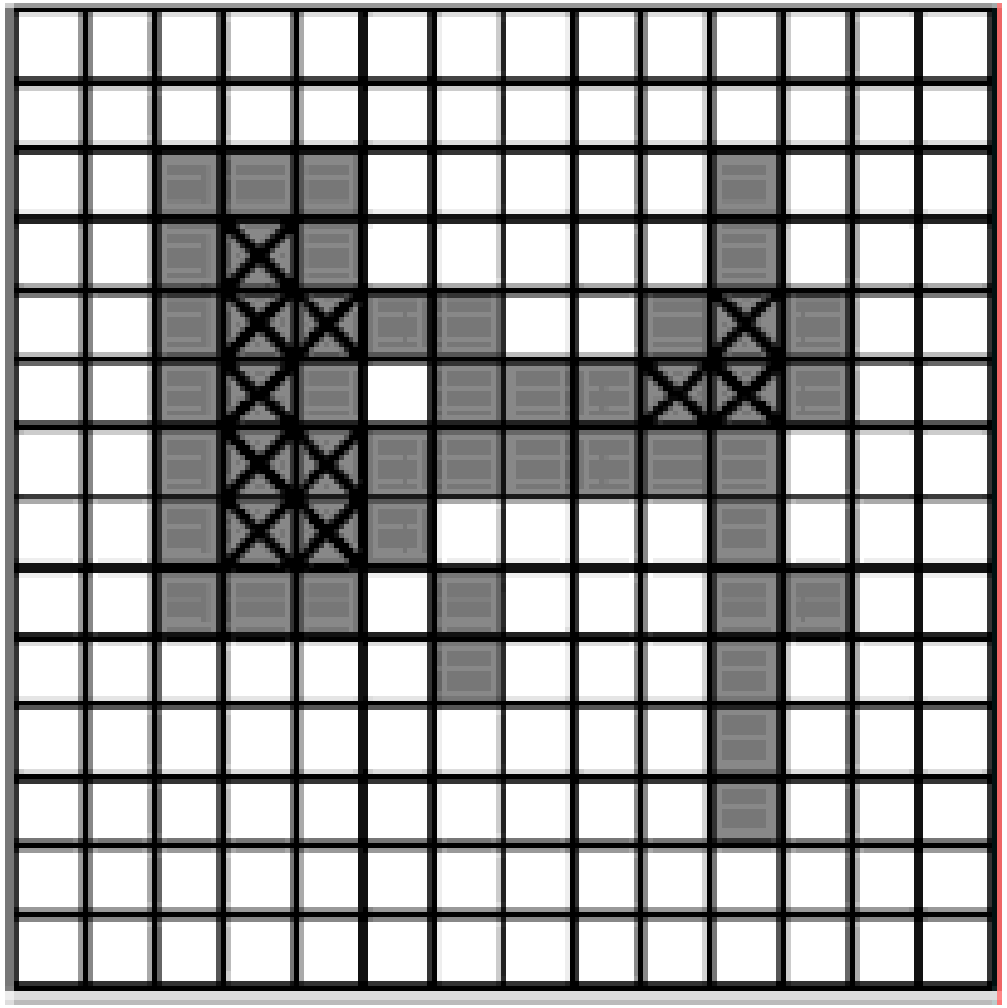
Erosion



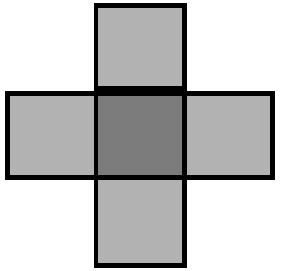
SE=



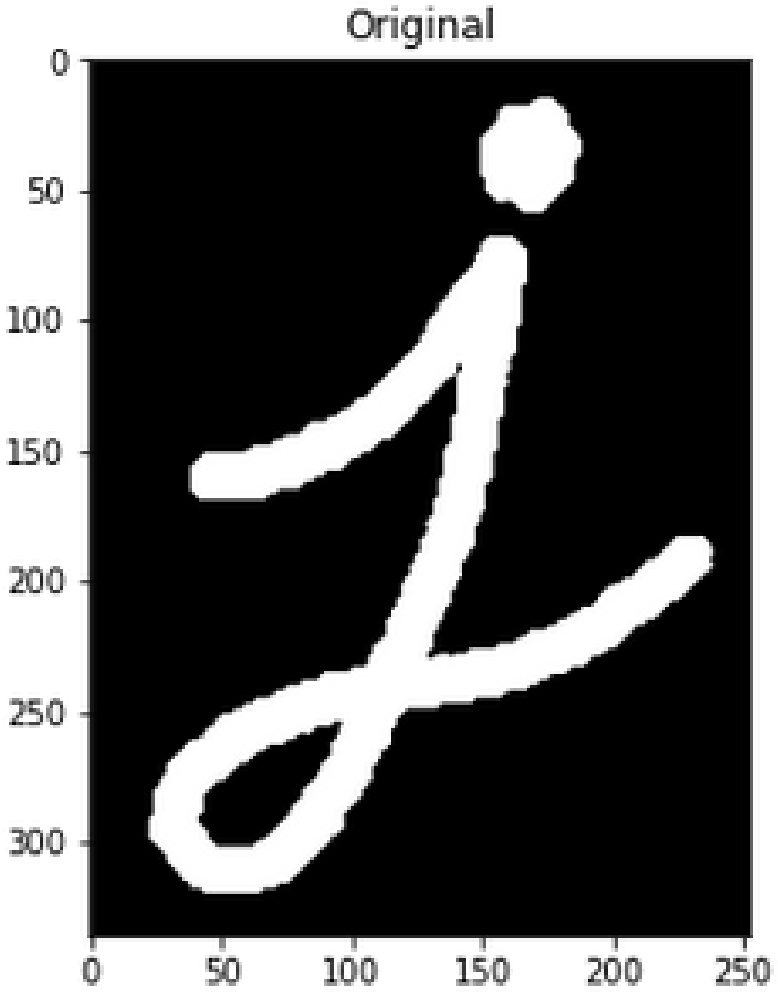
Erosion

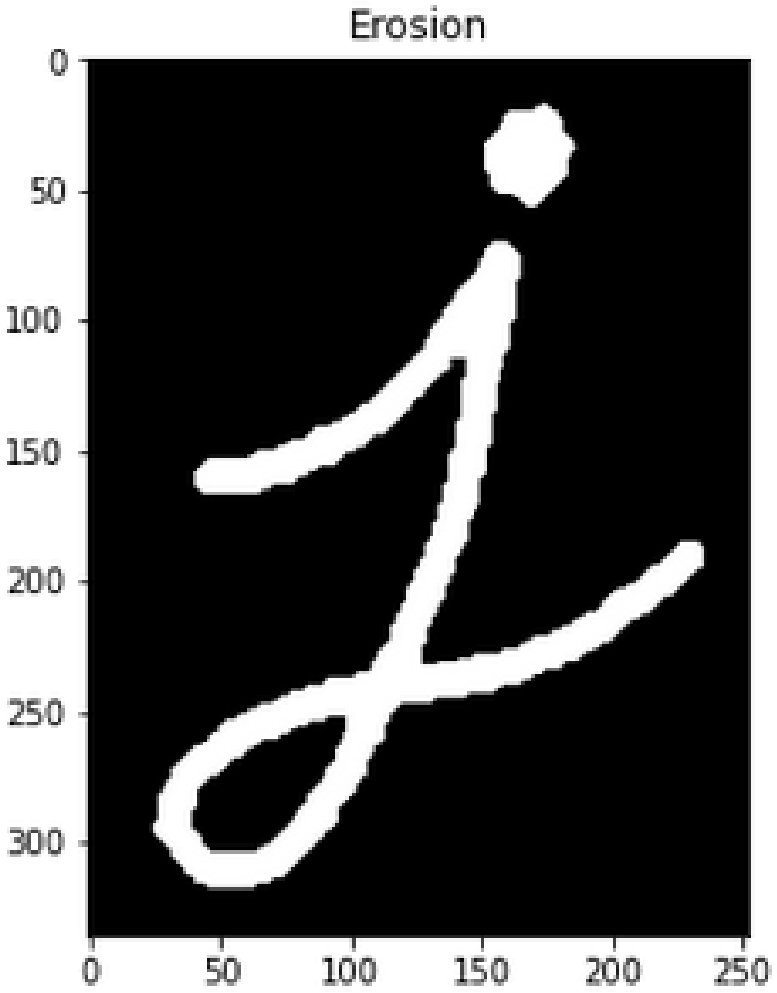


SE=

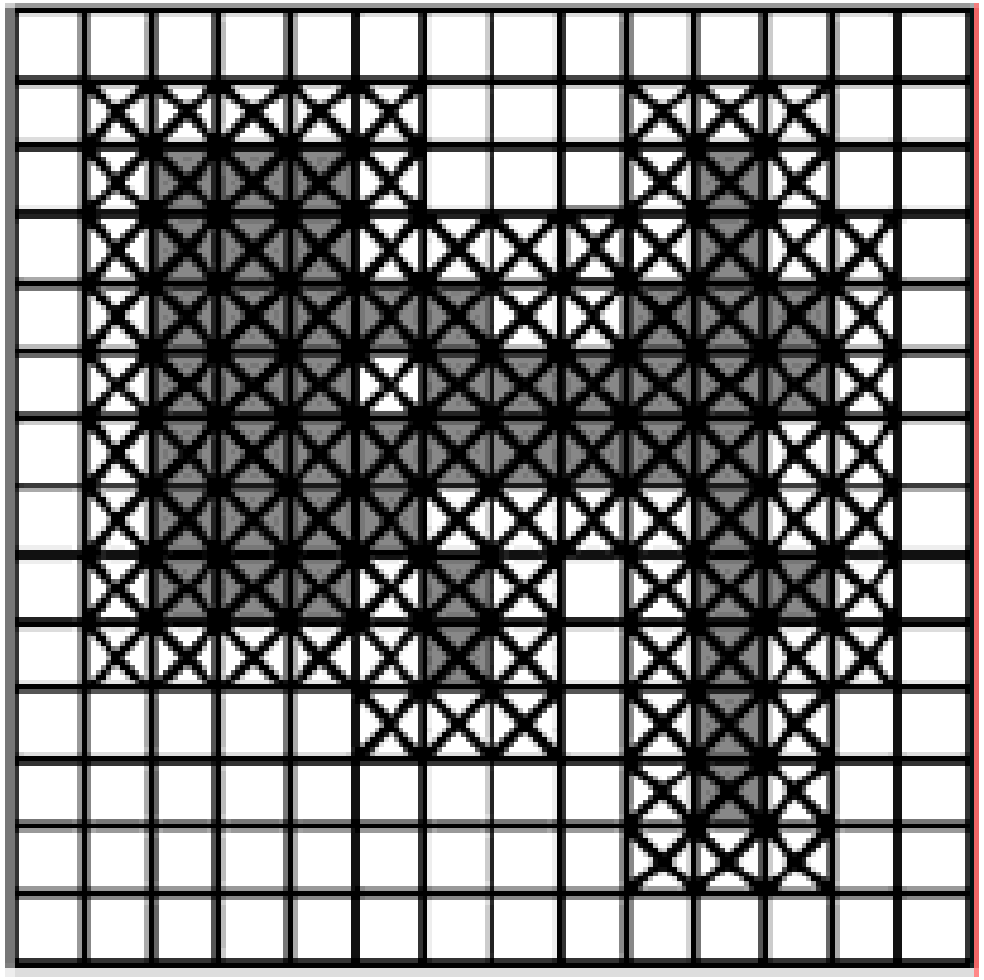


Erosion

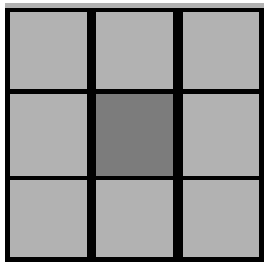

Original
0
50
100
150
200
250
300
0
50
100
150
200
250


Erosion
0
50
100
150
200
250
300
0
50
100
150
200
250

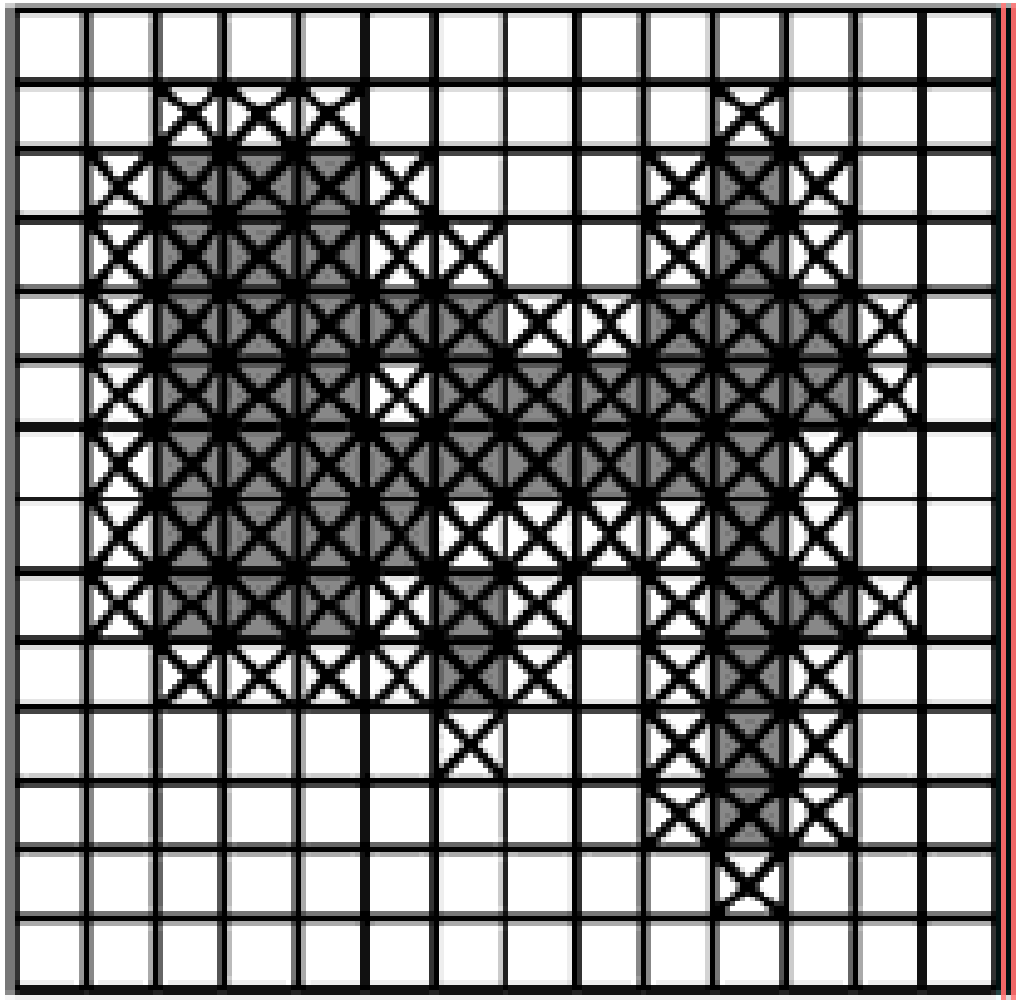
Dilation



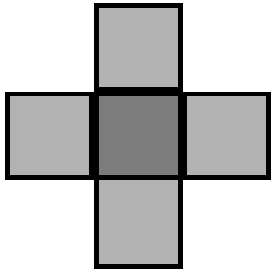
SE=



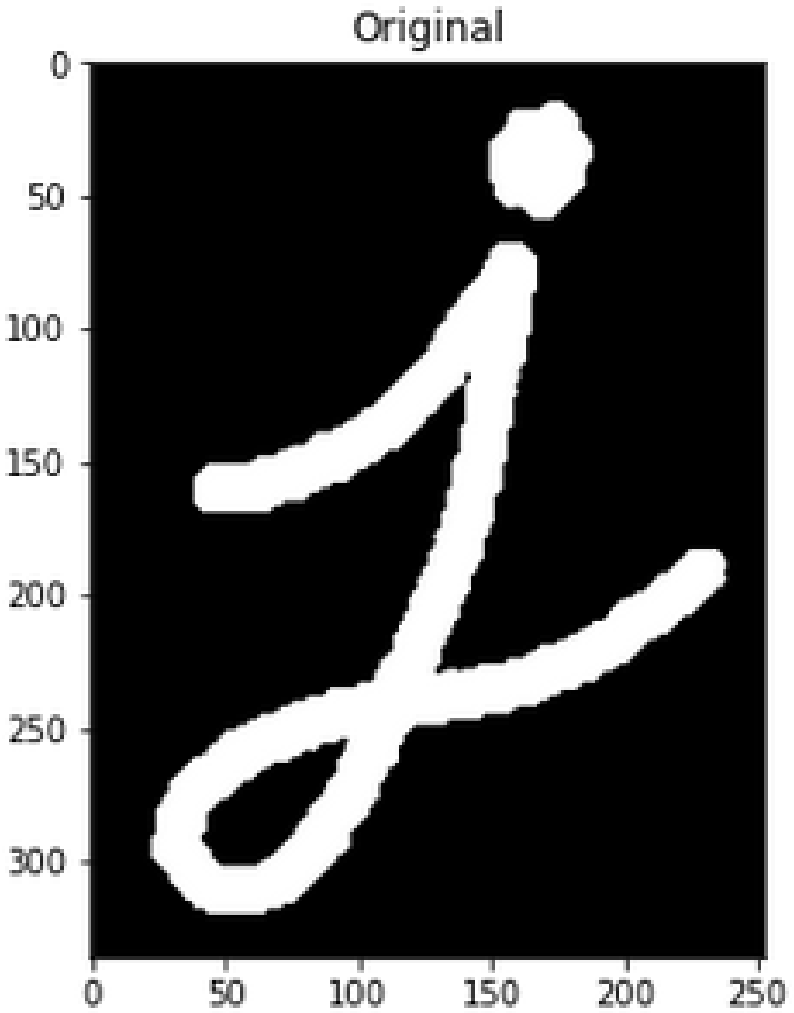
Dilation

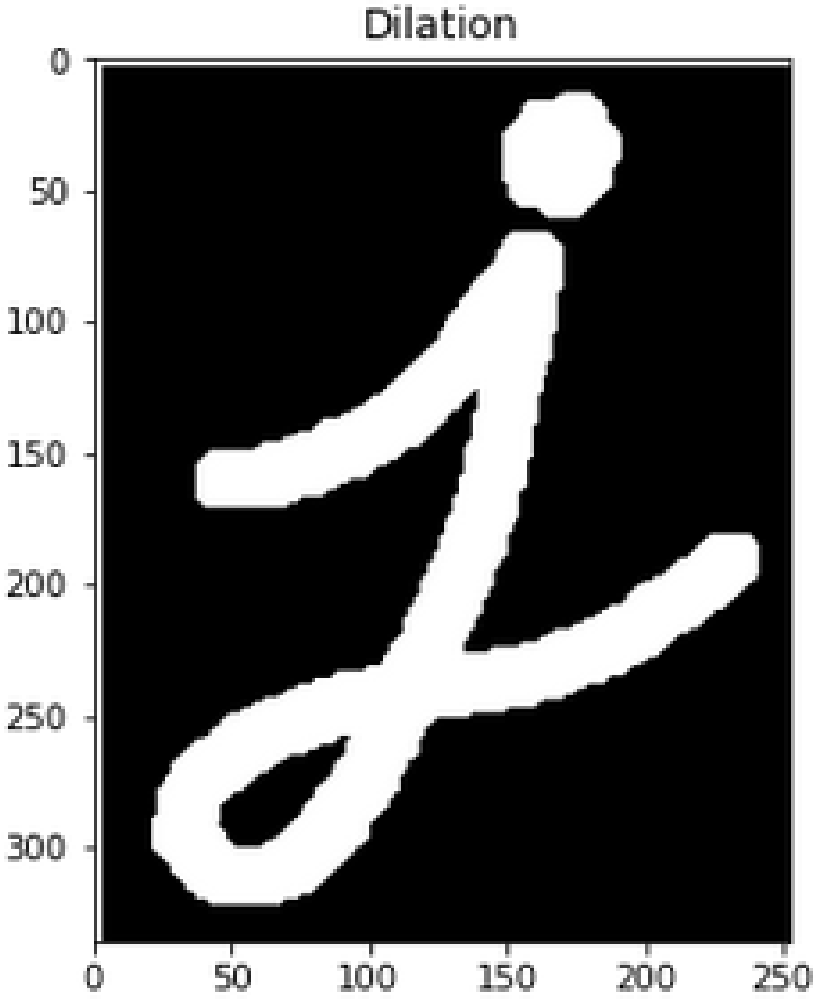


SE=



Dilation

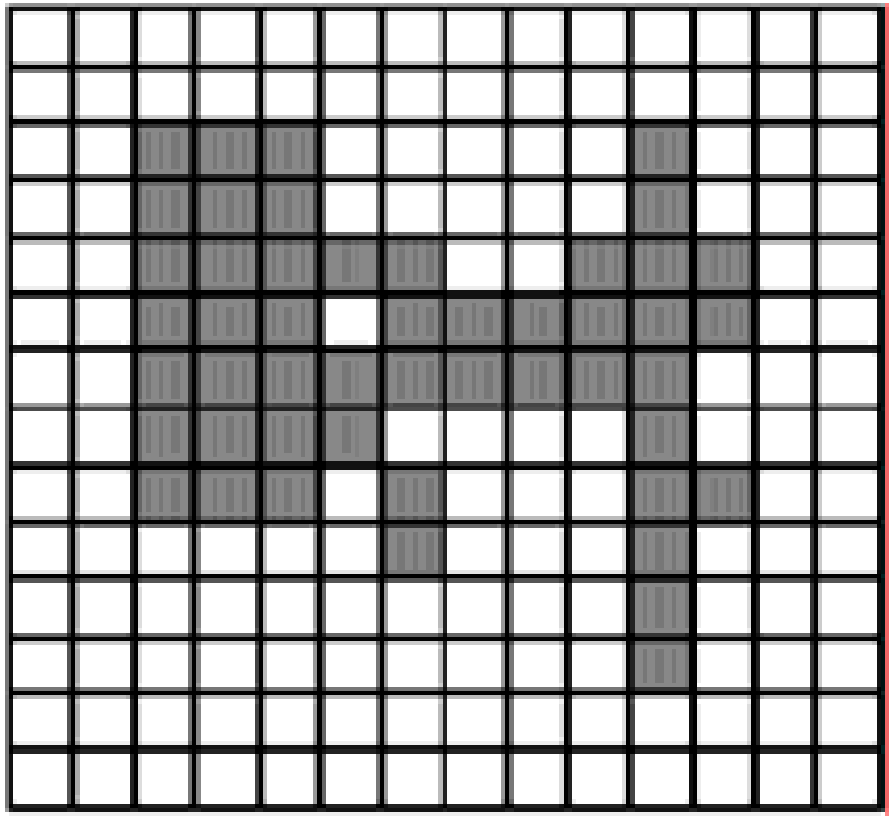

Original
0
50
100
150
200
250
300
0
50
100
150
200
250


Dilation
0
50
100
150
200
250
300
0
50
100
150
200
250

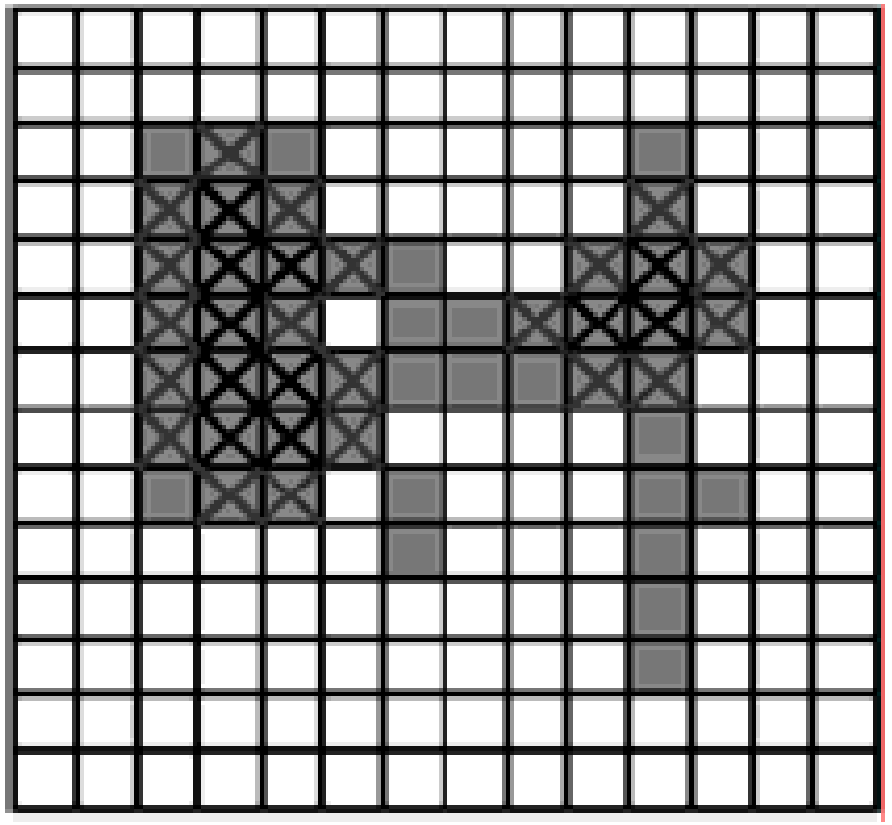
Opening

$$A \circ B = (A \ominus B) \oplus B$$

B=



A

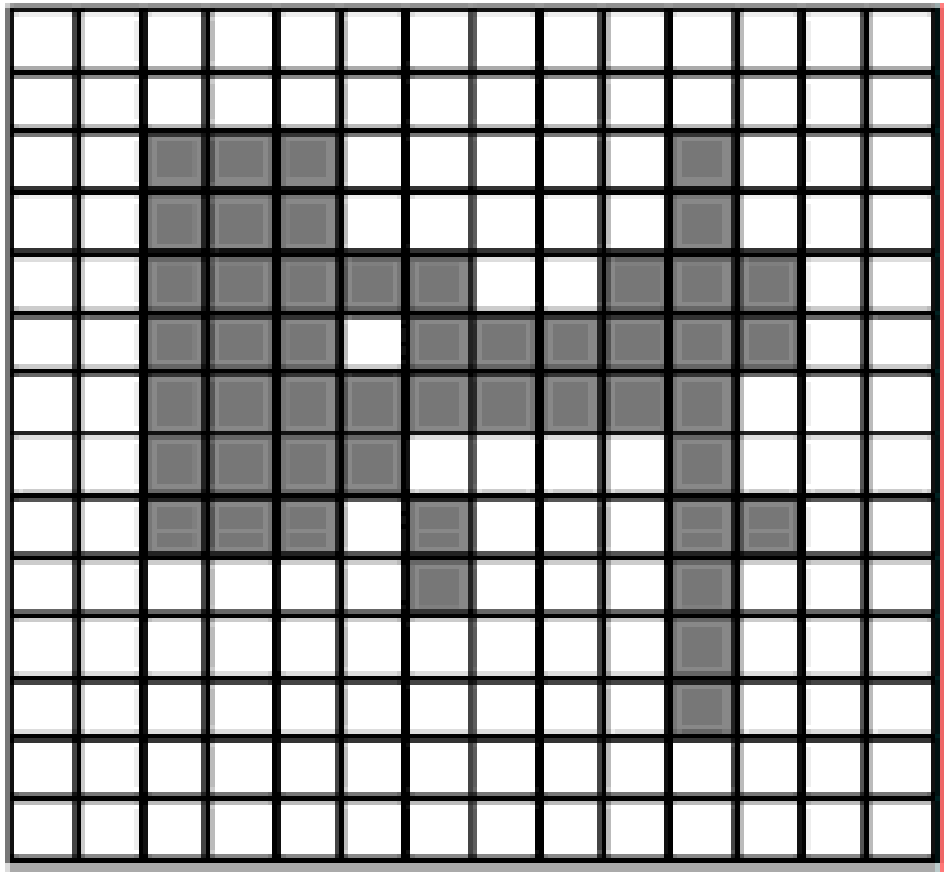


A⊖B A∘B

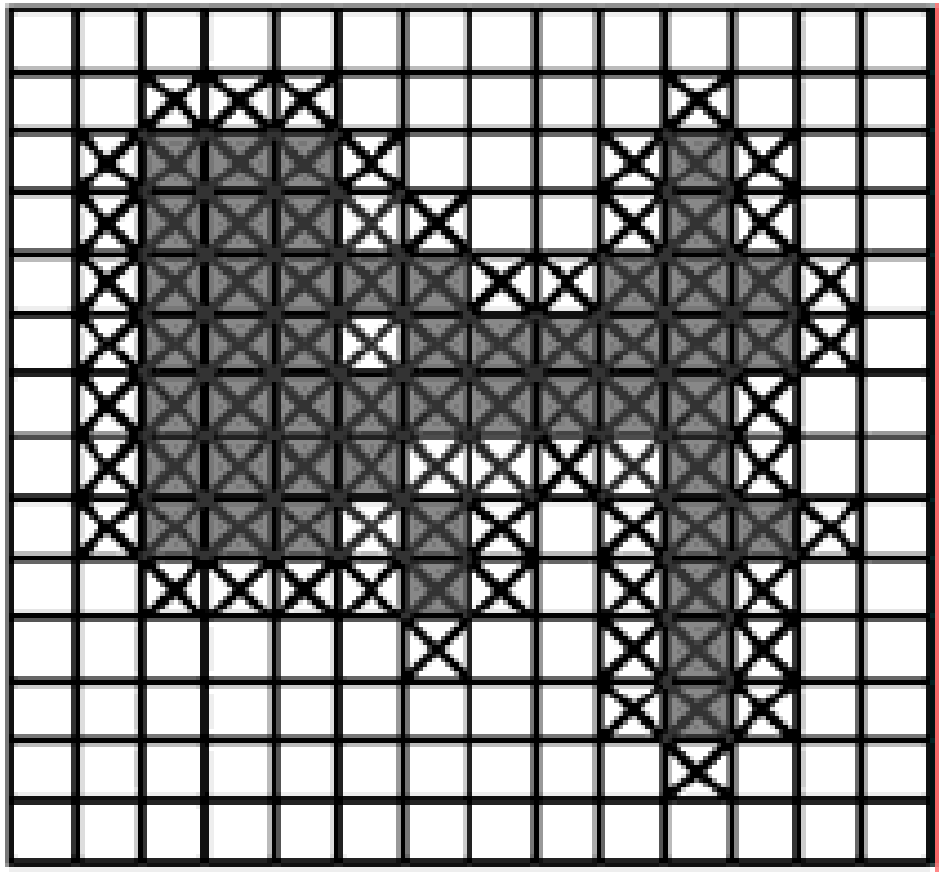
Closing

$$A \bullet B = (A \oplus B) \ominus B$$

B=



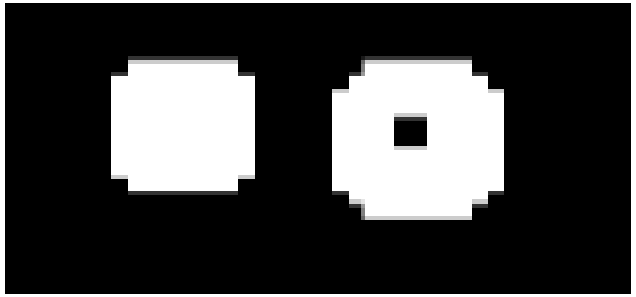
A



A⊕B A • B

Opening and Closing

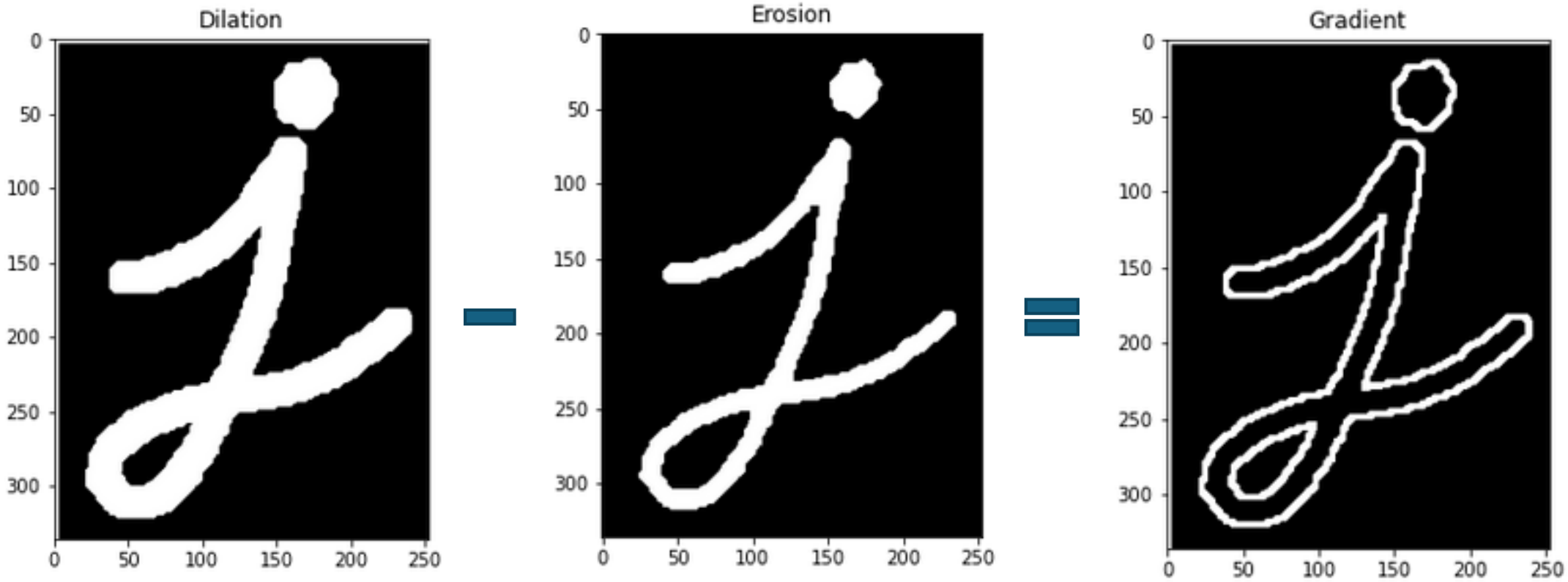
A



opening of A
→removal of small protrusions, thin connections, …

closing of A
→ removal of holes

Gradient



Thanks